



## **InfoRank: Information-Centric Autonomous Identification of Popular Smart Vehicles**

Junaid Ahmed Khan, Yacine Ghamri-Doudane, Dmitri Botvich

### **► To cite this version:**

Junaid Ahmed Khan, Yacine Ghamri-Doudane, Dmitri Botvich. InfoRank: Information-Centric Autonomous Identification of Popular Smart Vehicles. IEEE VTC Fall 2015, Sep 2015, Boston, MA, United States. pp.6, 10.1109/VTCFall.2015.7390985 . hal-01283729

**HAL Id: hal-01283729**

**<https://hal.science/hal-01283729>**

Submitted on 17 Mar 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# InfoRank: Information-Centric Autonomous Identification of Popular Smart Vehicles

Junaid Ahmed Khan<sup>\*†</sup> Yacine Ghamri-Doudane<sup>†</sup> and Dmitri Botvich<sup>\*†</sup>

<sup>\*</sup>University Paris-Est, LIGM Lab, Marne-la-Vallée, France

<sup>†</sup>L3i Lab, University of La Rochelle, France

khan@univ-mlv.fr, yacine-ghamri@univ-lr.fr, dbotvich@gmail.com

**Abstract**—Modern cars are transforming towards autonomous cars capable to make intelligent decisions to facilitate our travel comfort and safety. Such “Smart Vehicles” are equipped with various sensor platforms and cameras to collect, store and share tremendous amount of heterogeneous data from urban streets. This paper addresses the efficient collection and distribution of such massive data by allowing a popular Smart Vehicle to autonomously decide its user relevant importance in the vehicular network without relying on the infrastructure network. Therefore, we propose an Information-Centric algorithm, “InfoRank” for a vehicle to rank different location-dependent information associated to it. It then uses the information importance to analytically find its influence in the network. InfoRank is the first step towards identifying the best information hubs to be used in the network for the efficient collection, storage and distribution of urban sensory information. Results from scalable simulations using realistic vehicular mobility traces show that InfoRank is an efficient ranking algorithm to find top information facilitator vehicles in comparison to other ranking metrics in the literature.

**Keywords**—Information-Centric Vehicular Networking, Urban Sensing and monitoring

## I. INTRODUCTION

Autonomous cars are around the corner since vehicles, nowadays are equipped with a plethora of electronic components, sensors, cameras and wireless communication devices to assist drivers regarding the travel safety and comfort. They can be considered as an instance of the Internet of Things (IoT) centered at “Smart Vehicle” to collect and share different sensory and multimedia data from urban streets to offer various Intelligent Transportation System (ITS) applications in a Smart City. These include efficient traffic management, urban sensing and vicinity monitoring [1].

The key challenge is the efficient collection, distribution and storage of such massive amount of data with the intermittent connectivity and the vehicles mobility. Most of the content is of “local relevance” as the intended users lies within the vehicular network. Relying on the infrastructure network for the collection, storage and distribution of such heterogeneous *Big-Data* from vehicles can thus prove costly and inadequate to its usage. Pre-advertising or broadcasting all the sensing data from each vehicle would result in a massive advertising overhead and a redundant information storm within the network.

To address this issue, we propose to use popular Information Facilitator Vehicles (IFVs) with learning capabilities to anticipate the user interests within the citywide VANET. Selective IFVs are responsible for the efficient gathering,

storing and publishing of urban sensing data from source vehicles and deliver it to end users upon request. Therefore, for the first time in vehicular networks, we envision smart vehicles capable of finding their relative importance in the network as potential IFVs. Thus, the target of this paper is to introduce a ranking algorithm facilitating the identification of IFVs. First, the vehicle ranks the information associated to it taking into consideration the relevance to the users interest. It then considers the associated location-relevant information popularity to find its relative influence in the network using InfoRank algorithm as the *Vehicle Centrality*.

Our ranking algorithm considers Information-Centric Networking (ICN) [2] as the used communication paradigm. ICN is a content-centric networking architecture proposed to replace the current IP based Internet. In ICN, a user broadcasts an interest for content by its name, any corresponding host in the network replies back with the desired content. ICN aims to decouple the service from the host, thus removing content association to any physical location. Additionally, it offers In-Network caching at intermediate nodes while forwarding and responding to subsequent user interests. Our approach assumes ICN in VANETs as we consider this as a viable candidate in order to cater with the high mobility and intermittent connectivity of vehicles. The major contributions to this paper can be summarized as follows:

- We propose a novel distributed algorithm enabling a vehicle to rank important location-dependent information associated to it based on the satisfied user interests.
- The first vehicle ranking algorithm, “InfoRank”, is proposed, where each vehicle can find its influence in the network, without relying on any infrastructure network.
- We validate the scalability as well as ICN compliance by performing extensive simulations comprising around three thousand vehicles using realistic mobility traces to identify popular IFVs in time evolving VANETs.

The obtained results show that the proposed algorithms are well suited to help in the efficient identification of the top IFVs in the network using information-centric vehicular networking.

The rest of the paper is organized as follows. The next Section highlights the major related work. In Section III, we propose InfoRank followed by the performance evaluation discussing the simulations in Section IV. Section V concludes the paper along some insight for future research.

## II. RELATED WORK

Urban Sensing and vicinity monitoring using vehicles has attracted lots of researchers in the past few years and several schemes are proposed. One such scheme is Mobeyes [1] where sensor-equipped vehicles monitor the surrounding and relay aggregated summaries to mobile agents. Similarly [3] proposed a compressive sensing based monitoring in a delay tolerant vehicular network. In [4], real time multimedia data from urban streets is collected using vehicles. CarSpeak [5] allows vehicle to collaborate and access sensory information captured by neighboring vehicles in the same manner as it can access its own. Recently, we observe a shift towards ICN in [6], [7] and [8] as the underlying routing protocol for Vehicular Networks.

Centrality measures such as Degree, Closeness, Betweenness and Eigenvector centrality are used to find important nodes in the complex networks. Degree centrality considers the number of direct (one hop) neighbors of a node. Closeness centrality is the inverse of the sum of the lengths of the shortest paths from a node to the rest of the nodes in the network. Betweenness centrality is the fraction of all pairs of shortest paths passing through a node, where Eigenvector centrality the node's influence measure in the network [9]. Typical applications include social networks to identify influential information hubs for publishing/spreading information. Another interesting application is found in medical sciences to find epidemic disease spreaders [10]. Similarly, Google's PageRank [11] algorithm ranks the importance of a web-page in an Internet search based on the number of directed links towards it. In Delay Tolerant Networks (DTNs), BubbleRap [12] is an example where nodes with high centrality score are used for data dissemination.

Unlike social networks and Internet, it is unfeasible to use centrality-based popularity schemes in VANETs for multiple reasons; First, The rapid topological changes due to the high mobility of vehicles requires a continuous time varying analysis of the VANETs which is unfeasible by a practical scheme. Typical schemes assume a static graph topology with respect to time where the temporal network characteristics of VANETs would be ignored. Second, centrality measures such as Betweenness, Closeness and Eigenvector centrality computation requires network wide parameters, while in VANETs a vehicle cannot have such information to make run-time decisions. Third, existing schemes consider shortest path metric to compute a node's importance, while the highly dynamic VANET topologies does not ensure the availability of a stable path between nodes. Therefore, a new vehicle ranking algorithm adapted to VANETs and enabling vehicles to decide their relative importance in the network by overcoming the above mentioned constraints need to be thought about.

## III. INFORANK: A FULLY DISTRIBUTED APPROACH

InfoRank is a centrality measure enabling each IFV to autonomously find its importance in the network independent of a centralized database. The rapid changes in the frequency and duration of vehicle contacts cannot be used to decide its importance in the time evolving vehicular network. Though, the user's interest satisfaction for content is considered as a key metric for a vehicle's importance as it regularly responds to user interests. Therefore, we introduce a new metric where

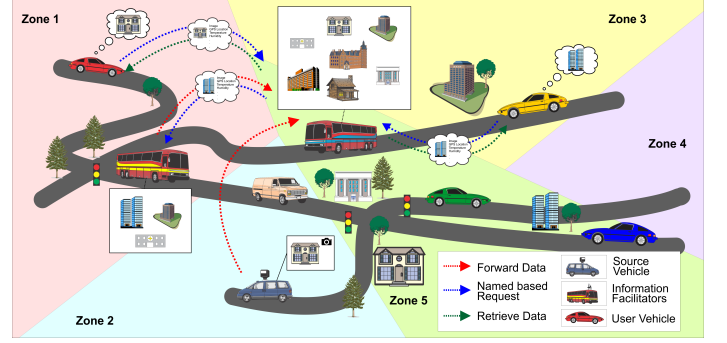


Figure 1: System Overview

each periodically finds its importance in the network with respect to the user relevant information. Figure 1 shows the final targeted system using selected IFVs for data collection, storage and distribution within VANET. The name-based interest is issued by user, which is forwarded to the nearby IFVs within VANET. The corresponding IFV replies with the desired content using the underlying ICN architecture. The following section defines the network model enabling each vehicle to compute its respective centrality.

### A. Network Model

We consider a time varying VANET modeled as an undirected vehicular graph  $G(\mathbb{V}(t), \mathbb{E}^v(t))$ , where  $\mathbb{V}(t) = \{v\}$  is a set of vertices  $v$ , each representing a vehicle on the road at time  $t$ .  $\mathbb{E}^v(t) = \{e_{jk}(t) \mid v_j, v_k \in \mathbb{V}, j \neq k\}$  is the set of edges  $e_{jk}(t)$  modeling the existence of a direct communication link between vehicles  $j$  and  $k$  at time  $t$ . The number of edges  $\mathbb{E}^v(t)$  depends on the transmission range of each vehicle. We assume it as a simple unit disk model bounded by the communication range  $R$ . The city map is represented by the undirected graph  $G(\mathbb{X}, \mathbb{E}^x)$ , the set of vertices  $\mathbb{X} = \{x\}$  contains different urban zones  $x$  and the set of edges  $\mathbb{E}^x = \{e_{pq} \mid x_p, x_q \in \mathbb{X}, p \neq q\}$  are their respective boundaries that connects different zones through a road network.

**Information Association:** Information association is defined as a bipartite graph  $G(\mathbb{V}, \mathbb{X}, \mathbb{E})$ , where  $\mathbb{V}$  is the set the vertices  $v$  in the vehicular graph  $G(\mathbb{V}(t), \mathbb{E}^v(t))$  and  $\mathbb{X}$  is the set of locations  $x$  in the city map  $G(\mathbb{X}, \mathbb{E}^x)$ . The edge  $\mathbb{E} = \{e_{ij} \mid v_i \in \mathbb{V}, x_j \in \mathbb{X}\}$  associates each vehicle to a set of regions  $X_v \subset \mathbb{X}$  with respect to the user relevant content.

The associated information is classified by clustering the regions using ICN hierarchical naming convention “/region/road-section/information-type”. Lane rules for road sections are usually defined by the city municipality while information type comprises different Intelligent ITS applications (Safety warnings, Road congestion information, Infotainment...) with varying content popularity and priority. The regions are clustered using voronoi tessellation [13] where the vehicles concentrated in an zone closer to a point are associated to the set of roads in a single voronoi region  $x \in G(\mathbb{X}, \mathbb{E}^x)$  as shown in Figure 1. For temporal VANET analysis, the time  $T$  is divided into set of regular time-slots  $\bar{t} = t_{k+1} - t_k$ , where each vehicle finds its centrality at the time instant  $t_{k+1}$  based on the known information in the current time-slot where  $t_k$  is the time instant at the beginning of the time-slot  $\bar{t}$ .

The information distance  $d(x, x_k)$  is the Euclidean norm between the content location  $x$  and  $x_k$ , the vehicles current position at time instant  $t_{k+1}$  where  $x, x_k \in X_v$ . We assume each vehicle knows the map of the city, i.e.  $G(\mathbb{X}, \mathbb{E}^x)$ . However, it only knows information relevant to itself, i.e. scope of the network information maintained in its cache. This is due to each vehicle's limited storage and coverage scope as it is unlikely for instance that it visited, and thus stored data, of all the traversable roads in the city. The content source could be either the vehicle itself (data sensed) or another vehicle for which it acted as an intermediate relay node. We will refer to content/information or location interchangeably in the text since content (information related to safety or traffic congestion) are associated to locations in the urban map.

### B. Information Importance

Information importance measures the vehicle relevance to users for a particular content. The interest-response frequency is a vital factor to classify content's importance. A vehicle holding popular content is considered as an important information hub in the network.

**Definition 1:** (Interest Satisfaction Frequency) We define  $I_x^v(\bar{t}) = \frac{r_x(\bar{t})}{R_x}$  as the frequency of user interests satisfied in the previous slot  $\bar{t}$ , where  $r_x(\bar{t})$  are the number of successful responds in the previous slot and  $R_x$  are the total successful responds for the content  $x \in X_v$  associated to the vehicle  $v \in \mathbb{V}$ .

IFVs regularly updates each content importance value depending on the interest satisfaction frequency. We assume that each IFV is capable to record the time and position each time it responds as the content provider to a user interest. Interest for each content specify the temporal scope of information validity, For instance, road congestion information is only valid during congestion. Therefore, in order to ensure the information importance is not substantially augmented after the desired deadline, let  $t_x^f$  be the last successful respond time for the content  $x$  and the average interest deadline as  $\bar{t}_d = \frac{1}{n} \sum_n t_x$  associated with each content, where  $n$  are the total number of interests in the previous time-slot and  $t_x$  is the deadline of each interest for content  $x$ .

**Definition 2:** (Information Timeliness) The information timeliness  $\tau(t_{k+1}) = \begin{cases} 1 & t_{k+1} \leq t_x^f + \bar{t}_d \\ e^{-\delta \bar{t}_d} & t_{k+1} > t_x^f + \bar{t}_d \end{cases}$  is the measure of the temporal information validity scope where  $\delta$  is the tuning parameter depending on the application needs (E.g. 1 hour for accident information validity).

If there are no active interests and the average interest validity time has passed, the information importance adapts an exponential delay since the information is of less importance in the network. However,  $\tau$  is set to unity for content to be always available in the network.

The corresponding content importance at the next time instant  $t_{k+1}$  is updated as follows:

$$C_x^v(t_{k+1}) = C_x^v(t_k) + \tau(t_{k+1}) I_x^v(\bar{t}) (1 + d(x, x_k))^{-\lambda} + s_x^v(t_{k+1}) \quad (1)$$

The content importance depends on  $C_x^v(t_k)$  at the beginning of the time-slot (time instant  $t_k$ ). If it is not responded in the

previous slot, then  $I_x^v(\bar{t}) = 0$  ensures the content importance is not increased unnecessarily. Here  $0 \leq s_x^v(t_{k+1}) \leq 1$  is the percentage of time the vehicle itself acted as the original source for any content  $x$ . The term  $s_x^v(t_{k+1})$  is updated regularly to ensure the content relevant to vehicle retain its value in case the vehicle does not respond in the previous slot. The interest later in time could finally route to the vehicle which maintains its value as the original source for particular content. The tuning parameter  $\lambda$  decide the value with respect to the associated content.

### C. Vehicle Centrality

The vehicle considers its importance with respect to the associated information in order to measure its influence in the networks. Besides information importance in Equation 1, we also consider the overall coverage scope as an important parameter to decide a vehicle importance in an urban environment.

**Definition 3:** (Coverage Entropy) We define  $H^v = - \sum_{\forall x \in \mathbb{X}} p(x) \log p(x)$ , as the coverage entropy of the vehicle periodically computed with respect to the entire city map (i.e. vehicle associated sub-graph  $X_v \in G(\mathbb{X}, \mathbb{E}^x)$ ). The probability  $p(x)$  is the visiting frequency to each region  $x \in \mathbb{X}$  before the importance computation time  $t_{k+1}$ .

Assume the vehicle's coverage in the map can be represented as a set of mobility between regions. The vehicles A and B coverage scope are bounded by the set of regions  $M^A = \{x_3, x_2, x_2, x_3, x_5, x_2\}$  and  $M^B = \{x_1, x_2, x_3, x_5, x_4, x_1\}$ . A visits the regions  $x_2, x_3$  and  $x_5$  with probabilities  $\frac{2}{6}, \frac{3}{6}$  and  $\frac{1}{6}$ , while B visits the regions  $x_1, x_2, x_3, x_4$ , and  $x_5$  with probabilities  $\frac{2}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}$  and  $\frac{1}{6}$  respectively. The corresponding coverage entropy is calculated as:

$$H^A = -\frac{2}{6} \log \frac{2}{6} - \frac{3}{6} \log \frac{3}{6} - \frac{1}{6} \log \frac{1}{6} = 0.439, \\ H^B = -\frac{2}{6} \log \frac{2}{6} - \left( \frac{1}{6} \log \frac{1}{6} * 4 \right) = 0.639,$$

Vehicle A has a narrow coverage scope due to its limited geographical coverage, while B has a wider geographical coverage with respect to the urban map. Therefore, we consider coverage entropy as the coverage metric for the vehicle importance with respect to all locations in the city.

Algorithm 1 shows the steps allowing a vehicle to find the respective InfoRank. For a given location-dependent content in cache, the corresponding information importance is updated for the next time-slot at time instant  $t_{k+1}$ . The information-centric centrality function is as follows:

$$f_I^v(t_{k+1}) = \frac{(1 + I_m^v(t_{k+1}))^{-\epsilon}}{|X_v|} \sum_{x \in X_v} C_x^v(t_{k+1}) \cdot w_x + H^v(t_{k+1}) \quad (2)$$

For all contents  $x \in X_v$  associated to  $v$ ,  $I_m^v(t_{k+1})$  are the ratio of missed interest to the total interests received by the vehicle while  $\epsilon$  is the tuning parameter. Missed interest provides the vehicle reliability regarding successful respond to the incoming interests.  $C_x^v(t_{k+1})$  is the respective content importance at time instant  $t_{k+1}$ ,  $w_x = \frac{R_x}{R_T}$  is the edge weight of information association graph  $G(\mathbb{V}, \mathbb{X}, \mathbb{E})$  considering the interest satisfied for the content  $x$  among all the contents in cache.  $R_x$  is the

number of responds for  $x$  and  $R_T$  is the number of responds for all contents in the cache.  $|X_v|$  is the cardinality of the sub-graph  $X_v \subset \mathbb{X}$ , all regions associated to the vehicle  $v \in \mathbb{V}$ .

The vehicle centrality at the time instant  $t_{k+1}$  is updated as the Exponential Weighted Moving Average (EWMA) function:

$$C_v(t_{k+1}) = (1 - \theta)C_v(t_k) + \theta f_I^v(t_{k+1})$$

where  $\theta$  is a tuning parameter to adjust the value for the past centrality score and the corresponding InfoRank in the current time-slot.

---

**Algorithm 1** InfoRank

---

**INPUT:**  $G(\mathbb{V}, \mathbb{X}, \mathbb{E})$  : information association graph  
**OUTPUT:** Updated InfoRank for the next time-slot at time-instant  $t_{k+1}$   
**for** each vehicle  $v \in \mathbb{V}$  **do**  
  **for** each content  $x \in X_v$  in cache **do**  
    Find  $d(x, x_k), \tau(t_{k+1}), s_x^v(t_{k+1}), w_x$   
    Compute  $I_x^v(\bar{t}) \leftarrow \frac{r_x(\bar{t})}{R_x}$   
    **if**  $I_x^v(\bar{t}) \neq 0$  **then**  
      Update  $C_x^v(t_{k+1})$  using Equation 1  
    **else**  
       $C_x^v(t_{k+1}) = C_x^v(t_k) + s_x^v(t_{k+1})$ ,  
    **end if**  
  **end for**  
  Find missed interests ratio  $I_m^v(t_{k+1})$ , Coverage entropy  $H^v(t_{k+1})$   
  Compute  $f_I^v(t_{k+1})$  using Equation 2  
**end for**  
**return**  $C_v(t_{k+1})$

---

#### IV. PERFORMANCE EVALUATION

One of the basic requirement for evaluating the efficiency of InfoRank is scalability. Therefore, we use Network Simulator-3 (NS-3)[14] as a scalable simulation platform for about three thousand vehicles. The performance of InfoRank is validated by a set of simulation runs under a realistic mobility scenario. We perform each simulation five times by analyzing different set of nodes as information producers and consumers in order to compute 95% confidence intervals. We rank the top information facilitators vehicles in the network by comparing their InfoRank score with the respective Degree, Closeness, Betweenness and Eigenvector centrality score.

Moreover, we consider a realistic mobility traces from Cologne, Germany to evaluate our proposed algorithms. To the best of our knowledge, it is considered as the most accurate mobility trace available for Vehicular Networks [15]. Starting from this trace, we regenerate a more easily usable version of it considering an area of  $6X6km^2$  for Cologne city center for 1-hour duration with a granularity of one second. The simulation parameters are summarized in Table I, followed by a description of the simulation scenarios used for the performance evaluation.

##### A. Simulation Scenario

We use the ndnSIM [16] module available for NS-3 to integrate Named Data Networking (NDN). The simulation

Table I: Simulation Parameters

Parameter	Value
Simulation platform	NS-3
Number of nodes	2986
Mobility trace	Cologne, Germany
Area	$6X6km^2$ city center
Duration	1 hour
Communication range	100m
Packet size	1024 bytes
Time granularity	1 sec
Simulation Runs	5

Table II: InfoRank in different set of Simulations

Simulation	1		2		3		4		5	
	ID	Score	ID	Score	ID	Score	ID	Score	ID	Score
1	34	1	1013	1	59	1	1210	1	10	1
2	2414	0.9719	1051	0.9966	63	0.9984	113	0.9930	56	0.9957
3	1239	0.9639	543	0.9958	108	0.9969	14	0.9923	1195	0.9838
4	108	0.9612	442	0.9919	348	0.9967	2372	0.9883	127	0.9833
5	517	0.9567	2366	0.9808	392	0.9958	184	0.9881	577	0.9801
6	414	0.9557	318	0.9788	152	0.9917	157	0.9875	658	0.9754
7	31	0.9516	2728	0.9770	48	0.9890	189	0.9834	5	0.9749
8	571	0.9506	268	0.9738	26	0.9839	197	0.9808	54	0.9742
9	502	0.9501	54	0.9735	37	0.9821	196	0.9803	1888	0.9718
10	2174	0.9489	150	0.9729	137	0.9819	30	0.9789	192	0.9714
										Mean
										1

scenario implements the following two applications:

**Producer:** Producer vehicle is the information source in the network where the areas visited in a time-slot before the InfoRank computation time are considered as content associated with the producer.

**Consumer:** Consumer vehicles are the potential user nodes planning to visit an area. Each consumer vehicle generates an interest for a content associated to a location in the city, which is routed to producer vehicles using the underlying ICN architecture.

It is crucial to choose a suitable scenario for analyzing the performance of InfoRank. For this reason, we consider urban sensing as our use case scenario as a special case of typical VANET operation. The city map is divided into voronoi zones/regions. The producers declare themselves as content source for the areas visited in the map. Consumers generate interests for the content associated with the area. We assume the interests follows a Zipf distribution, where interests for popular contents are more frequent [17]. This results in a high interests probability for popular locations in the city. The tuning parameters  $\lambda$ ,  $\delta$  and  $\epsilon$  are set to 0.5 in order to regulate the corresponding metrics.  $\bar{t}_d$  is set to 1 hour considering an urban sensing validity scope.

##### B. Simulation Results

The objective of our simulation study is to find answers to the fundamental question: *How well can it identify the top IFVs?* InfoRank score for the top 10 IFVs from five simulation runs are shown in Table II. For each rank, the average score lies within a confidence interval of 0.01 for a confidence level of 95%. Simulation traces for the top 100 nodes are obtained from five runs but here we show only the top ten nodes IDs and their score in each simulation for brevity. The centrality score is normalized with respect to the top identified node, i.e. the top node with a unity score followed by the relative score of other vehicles. We will use the same convention to interpret results in the later sections. In the first simulation, the vehicle 34 is identified to have the top InfoRank score among

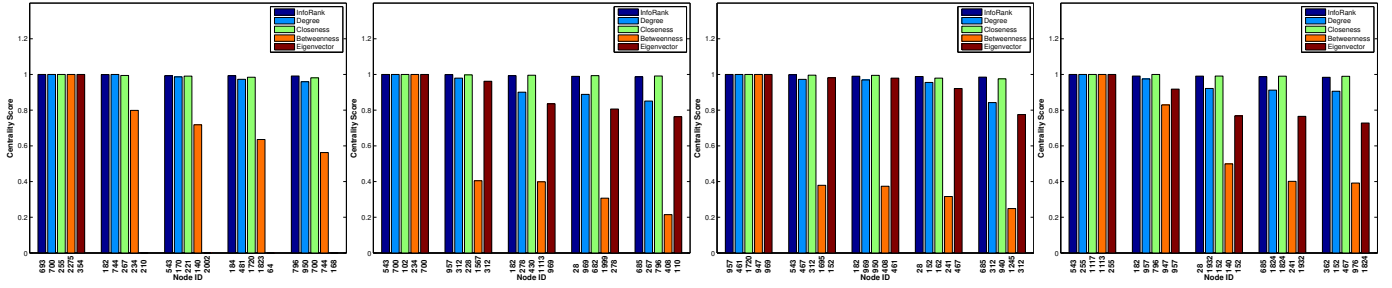


Figure 3: Temporal snapshots after each 15 minutes comparing top identified nodes by each schemes

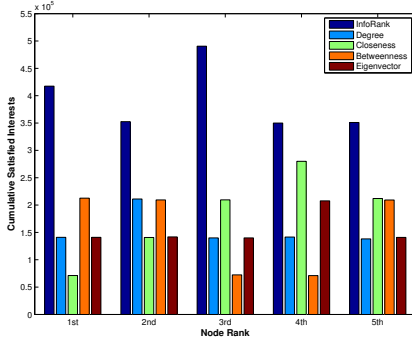


Figure 2: Cumulative Satisfied Interests by top identified nodes using each scheme over an average of five different simulation scenarios

the selected IFVs in the network as it satisfied the incoming interests more frequently throughout the simulation.

We consider the following performance metrics in comparison with the state of the art importance computation schemes (Degree, Closeness, Betweenness and Eigenvector centrality):

- Cumulative Satisfied Interests (CSI) for the top identified nodes by each scheme
- Comparison of top nodes identified by each scheme with their respective centrality scores
- Average aggregated throughput of the identified top ranked nodes by each scheme
- Cache hit rate for the top nodes by each scheme to evaluate InfoRank along ICN in VANET mobility scenarios

1) *Cumulative Satisfied Interests*: Cumulative Satisfied Interests refers to the total number of user interests satisfied after one hour. Figure 2 shows the CSI score of the top five nodes identified by all these schemes in an average of five set of simulations. Typical ranking schemes only takes into account physical topology towards computing a node importance in the network, ignoring the satisfied user interests. Nevertheless, InfoRank satisfied more user interests than other schemes in all the five set of simulations due to the consideration of user interest satisfaction as a key factor towards a vehicle importance in the network.

2) *Temporal behavior analysis of top nodes*: It is important to efficiently analyze the time varying behavior of

our algorithm due to the dynamic VANET environment. The time varying behavior of the relative score of the top five nodes identified by all schemes are shown by periodic network snapshots after each 15 minutes interval in Figure 3. We consider the top node identified by each scheme as benchmark by assigning it a unity score. At the beginning, vehicle 693 is ranked as top IFVs by InfoRank, though the other schemes underrated it. Vehicles also change places along the ranking order. For example, around 30 minutes, the node 543 replaced 693 as the top IFV, then it is replaced by the node 975 around 45 minutes and finally it regained the top position.

An interesting results was observed around 15 minutes: Only one node yields a high Eigenvector centrality score. We investigate this effect and found that the principle eigenvalue yields the top node where the eigenvector is shifted towards the principle component. Thus, resulting in one major central node. This shows that the famous Eigenvector centrality fails to assign significant score to a large fraction of nodes in a large network, while InfoRank do not reflect such behavior. Other centrality schemes result in different set of top nodes at every snapshot since such schemes only consider the instantaneous shortest paths requiring complete topological information. However, such complete network information is not available to an individual vehicle in highly unstable VANETs. InfoRank is not affected by such network dynamics since we are able to rank each vehicle considering relatively stable metrics.

3) *Aggregated Per Node Throughput*: We also evaluate the ranking scheme by analyzing the throughput at important nodes in the network. Figure 4 shows the aggregated per node throughput of the top nodes identified by each scheme. The average aggregated throughput (Kbps) is computed over the entire simulation duration for five set of simulations. The top nodes identified by InfoRank yields more throughput compared to other schemes. We observe that the throughput of the third node is relatively higher, thus inferring a variation between different ranks. Similar variation is seen for Degree, Betweenness and Eigenvector centrality. However, InfoRank outperformed all schemes as it incorporates information importance factor towards vehicle importance computation, while other schemes rely only on topological measures (node degree or shortest paths) towards vehicle importance computation.

4) *ICN Evaluation - In-Network Caching*: We evaluate the ICN built-in feature of In-Network caching at the intermediate nodes by computing the cache hit rate at the top nodes identified by each scheme as shown in Figure 5. A second successful response by a node for the same content is consid-



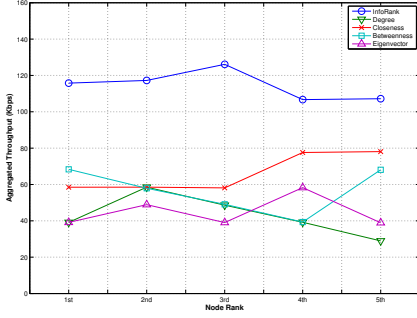


Figure 4: Average aggregated throughput by the top identified nodes using each scheme in five simulations

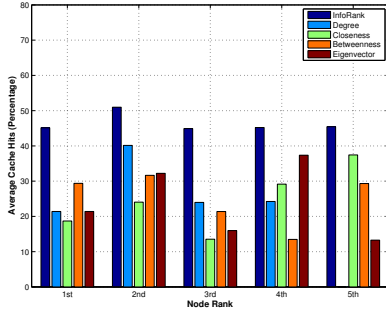


Figure 5: Average cumulative cache hit rate by the top identified nodes using each scheme in five simulations

ered a cache hit. The cumulative cache hit rate is computed for the entire simulation duration for five set of simulations. The top nodes identified by InfoRank yield a higher hit rate than all the other schemes in all the five simulations. This is because InfoRank considers information importance as a key factor, thus, the vehicle containing important information responds and subsequently cache more frequently compared to other vehicles. This proves that In-Network caching offered by ICN in InfoRank implementation overcomes the mobility and intermittent connectivity constraints in VANETs for efficient content access.

Responding to the question posed before, *How well can it identify the top IFVs?* From the simulation results, It is clear that a relatively stable set of top IFVs are identified by InfoRank compared to the other schemes in dynamic VANETs. It is also shown that InfoRank can identify nodes which satisfied more user interests with higher aggregated per node throughput and more cache hit rate compared to the other schemes. Thus, the overall comparative analysis of InfoRank with different network ranking schemes in the literature proved it as an efficient vehicle ranking algorithm.

## V. CONCLUSIONS AND FUTURE DIRECTIONS

We introduced the first vehicle ranking algorithm “InfoRank”, enabling popular Smart vehicles to rank themselves in a fully distributed VANET. InfoRank first ranks important location-dependent information based on the user interests satisfaction frequency. We then employed InfoRank using

content-centric networking architecture supporting in-network caching to cater high mobility and intermittent connectivity in VANETs. Extensive simulations were performed using realistic mobility traces to evaluate the efficiency of the proposed algorithm and a time varying network behavior is analyzed. Results by comparing with state of the art centrality schemes revealed that InfoRank is best suited to efficiently identify important information facilitator vehicles in VANETs.

Identification of top IFV can be helpful in different smart city applications such as urban sensing mentioned in the paper. Popular IFVs can be used for efficient data collection, storage and distribution in VANETs. Designing such efficient schemes will be the subject of our future research.

## ACKNOWLEDGMENT

This work is funded by the French Ministry of Industry under the CarCoDe (ITEA2 11037) project.

## REFERENCES

- [1] U. Lee, B. Zhou, M. Gerla, E. Magistretti, P. Bellavista, and A. Corradi, “Mobeyes: smart mobs for urban monitoring with a vehicular sensor network,” *Wireless Communications, IEEE*, vol. 13, no. 5, pp. 52–57, 2006.
- [2] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, and B. Ohlman, “A survey of information-centric networking,” *Communications Magazine, IEEE*, vol. 50, no. 7, pp. 26–36, 2012.
- [3] H. Wang, Y. Zhu, and Q. Zhang, “Compressive sensing based monitoring with vehicular networks,” in *INFOCOM, 2013 Proceedings IEEE*. IEEE, 2013, pp. 2823–2831.
- [4] R. Bruno and M. Nurchis, “Efficient data collection in multimedia vehicular sensing platforms,” *Pervasive and Mobile Computing*, 2014.
- [5] S. Kumar, L. Shi, N. Ahmed, S. Gil, D. Katabi, and D. Rus, “Carspeak: a content-centric network for autonomous driving,” *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 4, pp. 259–270, 2012.
- [6] P. Talebifard and V. Leung, “Towards a content-centric approach to crowd-sensing in vehicular clouds,” *Journal of Systems Architecture*, vol. 59, no. 10, pp. 976–984, 2013.
- [7] M. Amadeo, C. Campolo, A. Molinaro, and G. Ruggeri, “Content-centric wireless networking: A survey,” *Computer Networks*, vol. 72, pp. 1–13, 2014.
- [8] M. Gerla, E.-K. Lee, G. Pau, and U. Lee, “Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds,” in *Internet of Things (WF-IoT), 2014 IEEE World Forum on*. IEEE, 2014, pp. 241–246.
- [9] S. P. Borgatti, “Centrality and network flow,” *Social networks*, vol. 27, no. 1, pp. 55–71, 2005.
- [10] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, “Identification of influential spreaders in complex networks,” *Nature Physics*, vol. 6, no. 11, pp. 888–893, 2010.
- [11] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation ranking: Bringing order to the web,” 1999.
- [12] P. Hui, J. Crowcroft, and E. Yoneki, “Bubble rap: Social-based forwarding in delay-tolerant networks,” *Mobile Computing, IEEE Transactions on*, vol. 10, no. 11, pp. 1576–1589, 2011.
- [13] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*. John Wiley & Sons, 2009, vol. 501.
- [14] “Network simulator 3 - <http://www.nsnam.org>.”
- [15] S. Uppoor, O. Trullols-Cruces, M. Fiore, and J. M. Barcelo-Ordinas, “Generation and analysis of a large-scale urban vehicular mobility dataset,” *Mobile Computing, IEEE Transactions on*, vol. 13, no. 5, pp. 1061–1075, 2014.
- [16] A. Afanasyev, I. Moiseenko, and L. Zhang, “ndnSIM: NDN simulator for NS-3,” NDN, Technical Report NDN-0005, October 2012. [Online]. Available: <http://named-data.net/techreports.html>
- [17] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, “Web caching and zipf-like distributions: Evidence and implications,” in *INFOCOM’99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 1. IEEE, 1999, pp. 126–134.